# Digital Lighthouse: A Platform for Monitoring Public Groups in WhatsApp

Ivandro Claudino de Sá[1], José Maria Monteiro[1], José Wellington Franco da Silva[1], Leonardo Monteiro Medeiros[1], Pedro Jorge Chaves Mourão[2] and Lucas Cabral Carneiro da Cunha[1]

[1]*Computer Science Department, Federal University of Ceará, Fortaleza, Ceará, Brazil*

[2]*Department of Sociology, Ceará State University, Fortaleza, Ceará, Brazil*

Keywords: Misinformation Detection, Natural Language Processing, WhatsApp, Social Media.

Abstract: The large-scale dissemination of misinformation through social media has become a critical issue, harming social stability, democracy, and public health. In Brazil, 48% of the population uses WhatsApp to get news. So, many groups have been used this instant messaging application to spread misinformation, especially as part of articulated political or ideological campaigns. In this context, WhatsApp provides an important feature: the public groups. These groups are so suitable for misinformation dissemination. Thus, developing software frameworks to monitor the misinformation spreading in WhatsApp public groups has become a field of high interest both in academia, government and industry. In this work, we present an entire platform, called Digital Lighthouse, that aims for finding WhatsApp public groups, besides extracting, cleaning, analyzing, and visualizing misinformation that spread in such groups. Using the Digital Lighthouse, we built three different datasets. We hope that our platform can help journalists and researchers to understand the misinformation propagation in Brazil.

## 1 INTRODUCTION

In the last years, the popularity of instant messaging applications has contributed to the spread of misinformation. Through these systems, misinformation can deceive thousands of people in a short time (due to their appealing nature) and cause significant harm to individuals or society. In this context, misinformation has been used to change political scenarios, to contribute to the spread of diseases, and even to cause deaths (Su et al., 2020).

The WhatsApp instant messaging application is very popular in Brazil, with more than 120 million users. In Brazil, 48% of the population use WhatsApp to get, share and discuss news. WhatsApp makes it possible to instantly share different media types, such as images, audios, and videos. Besides, WhatsApp provides a significant feature: the public groups. These public groups are accessible through invitation links published on popular websites and various social networks, such as Facebook and Twitter. Usually, they have specific topics for discussion, such as politics and education. In this way, WhatsApp public groups are very similar to social networks

Public groups have been used to spread misinformation, especially as part of articulated political or ideological campaigns. Furthermore, misinformation spreads faster, deeper, and expansive than legit information. Further, due to the high volume of information that we are exposed to, we have a limited ability to distinguish true information from misinformation (Vosoughi et al., 2018; Qiu et al., 2017).

In this context, monitoring the content that circulates in public WhatsApp groups is a fundamental task to understand the spread of misinformation and get insights to address this problem. However, collecting a database of WhatsApp messages is a challenging task. To fill this gap, we built the Digital Lighthouse, an entire platform that aims for finding WhatsApp public groups, besides extracting, cleaning, analyzing, and visualizing misinformation that spread in these groups. Early detection of misinformation could prevent its spread, thus reducing its damage. Using the Digital Lighthouse, we build three different WhatsApp' messages datasets, covering relevant themes such as the Brazilian general elections campaign in 2018, the covid-19 pandemic, and the vaccine for covid-19.



Claudino de Sá, I., Monteiro, J., Franco da Silva, J., Medeiros, L., Mourão, P. and Carneiro da Cunha, L.
Digital Lighthouse: A Platform for Monitoring Public Groups in WhatsApp.
DOI: 10.5220/0010480102970304
In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 1*, pages 297-304
ISBN: 978-989-758-509-8

The remainder of this paper is organized as follows. Section 2 presents the main related work. Section 3 describes the Digital Lighthouse platform. Section 4 details a case study performed to evaluate the proposed platform. Conclusions and future work are presented in Section 5.

## 2 RELATED WORK

It is essential to highlight that WhatsApp is unique in several ways relative to other social media platforms. WhatsApp was developed to allow users to privately send messages to each other through their smartphones. A specific aspect of WhatsApp messaging is the public groups. These are openly accessible groups, frequently publicized on well-known websites, and typically themed around particular topics. It is worth mentioning that texts extracted from WhatsApp are quite different from those collected through Websites, fact-checkers, or other kinds of social media platforms, such as Twitter. WhatsApp messages include conversation, opinions, humorous and satirical texts, prayers, commercial offers, news, short texts, emojis, and others.

Thus, despite the scientific community's efforts, there is still a need for monitoring and identifying misinformation in WhatsApp messages, mainly in Portuguese. The paper presented in (Garimella and Tyson, 2018) is a seminal work in collecting and analyzing WhatsApp messages. The authors built a dataset by crawling 178 public groups, containing 45K users and 454K messages, from different countries and languages, such as India, Pakistan, Russia, Brazil, and Colombia. In (Gaglani et al., 2020), the authors contextualize the problem of spreading fake news on WhatsApp, especially in India and Brazil, and proposes a strategy for the automatic detection of fake news. A total of 10 public groups were scraped for one week to get 1000 multilingual messages. In (Resende et al., 2018), the authors presented a system for gathering, analyzing, and visualize public groups in WhatsApp. Besides, the authors also provide a brief characterization of the 169.154 messages shared by 6,314 users in 127 public groups. In the study presented in (Machado et al., 2019), the authors collected and analyzed 298,892 WhatsApp messages, from 130 public groups, in the period of the 2018 Brazilian presidential elections. In (Resende et al., 2019), the authors analyzed different aspects of WhatsApp messages from public political-oriented groups. However, none of these works provides an entire public platform for finding, gathering, analyzing, and visualizing WhatsApp messages.

Other works propose classifiers to detect misinformation automatically (Silva et al., 2020; Faustini and Covões, 2019). In (Shu et al., 2018), the authors investigated the use of complex networks to detect and mitigate fake news on social media. During fake news dissemination, different entities can be categorized into content, social and temporal dimensions. These dimensions have mutual relations and dependencies. So, fake news dissemination has inherent network properties. In (Shu et al., 2019), the authors explored user profiles to detect fake news. They argue that there are correlations between malicious accounts and fake news. In this same way, the paper presented in (Hamdi et al., 2020) proposed a hybrid approach that explores features from the user profile and his social graph (Twitter followers/followees graph) to detect fake news. In (Zhang and Hara, 2020), the authors propose a probabilistic model for malicious user and rumor detection (MURD).

## 3 THE DIGITAL LIGHTHOUSE PLATFORM

This section will present the main components of the Digital Lighthouse platform, which aims to extract, analyze, and visualize misinformation in WhatsApp messages. The proposed platform architecture comprises four modules, as illustrated in Figure 1. The main contribution of this work is the orchestration of all these components, which will be detailed next.

### 3.1 Module I: Finding Public Groups

WhatsApp allows you to join public groups through the use of links (URLs) containing the domain 'chat.whatsapp.com' and a group identification code. These links are publicized through websites or social networks. In this way, groups can be found through queries on search engines like Google, or simply by accessing sites created for this specific purpose. This work used both strategies for finding public groups.

#### 3.1.1 Finding Web Pages with Invite Links

In order to find invitations links for WhatsApp public groups through the Google search engine, we develop a web crawler using the Python programming language. The crawler builds queries, sends them to the Google search engine and receives the result (links for web pages). To set up a particular query, the crawler receives a series of input parameters, sucha as: the WhatsApp domain, a set of keywords, and the target language. After a given query be executed, the

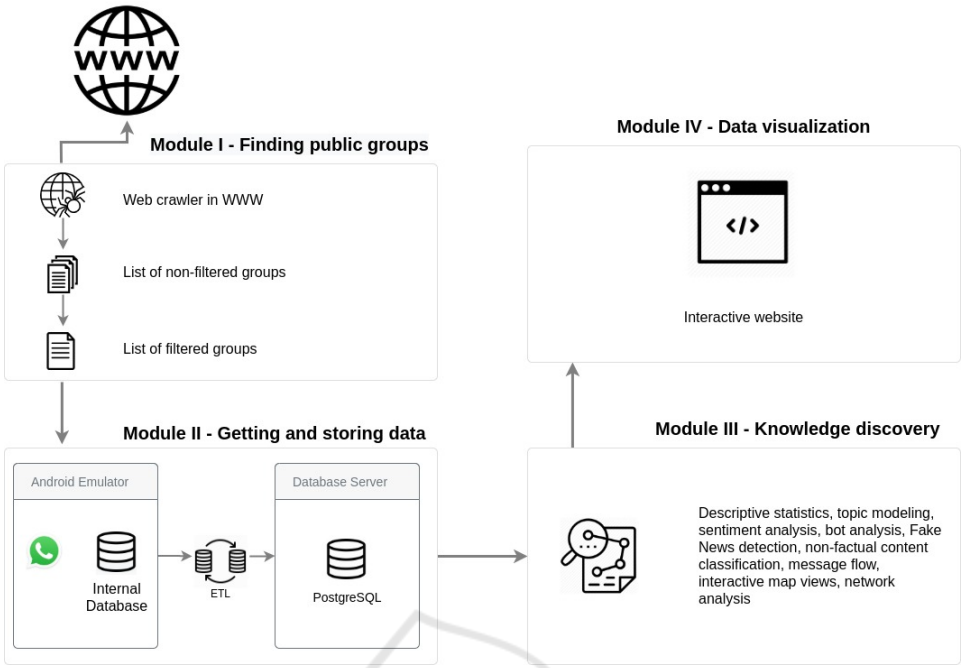


Figure 1: The Digital Lighthouse Platform Architecture.

crawler receives a set of metadata, including references to the web pages where the invite links were found. These web page links are stored in a file called search_links.csv.

#### 3.1.2 Finding Invite Links

The next step consists of requesting each web page found previously (and stored in the search_links.csv file) and parse it seeking for WhatsApp invite links. More specifically, the crawler sends a HTTP request for a certain web page link (URL). Then, the search engine answer the request by returning the web page content. After, a scraper will create a tree structure with the HTML content of the web page. This tree structure will be used to search for invite links. Finally, the scraper produces as output a list of invite links which are stored in a file called group_links.csv or yet list of non-filtered groups.

#### 3.1.3 Selecting Valid Invite Links

However, find a set of invite links is not sufficient. Some groups no longer exist, several links have been disabled, and a few groups have a tiny number of participants. Thus, it is necessary to check the status of each invite link. After this checking process, a new file, called a list of filtered groups, is generated containing only the valid links.

#### 3.1.4 Joining Public Groups

Finally, with valid links, it is possible to join public groups using a cell phone chip and a web browser, in an automatic or manual manner. In this work, we manually joined the groups to don't violate WhatsApp politics.

### 3.2 Module II: Getting and Storing Data

Unlike other social media, such as Twitter and Facebook, and due to its private chat nature, there is no public API to collect data from WhatsApp in an automated manner. Thus, monitoring WhatsApp public groups poses a technical and even ethical challenge. To tackle this issue, we take an approach similar to (Garimella and Tyson, 2018; Resende et al., 2018). Thus, in order to automatically collect the content (messages, audio, images, and videos) of the public groups that Digital Lighthouse joined, it have used WhatsApp Web and Selenium Web Driver.

#### 3.2.1 Getting the Content of Public Groups

The Digital Lighthouse uses a virtual machine (VM) containing an Android emulator, the WhatsApp Web, the Selenium Web Driver and a PostgreSQL database server. In the Android emulator we had installed the

WhatsApp application and a SQLite database. Finally, we used the Selenium Web Driver to manipulate the Android emulator and the WhatsApp Web in order to automatically ccess the public groups content and store it in the SqlLite database.

#### 3.2.2 Storing the Content of Public Groups

The messages extracted from WhatsApp are stored, in their original format, in a SQLite database. However, for that such messages can be effectively used for the purpose of knowledge discovery or to get insights, it is necessary that they undergo to a process of cleaning, integration and anonymization. After this process, the treated messages are stored in a PostgreSQL database, and can now be used for analysis and visualization purposes. It is important to highlight that the audios, images, and videos are stored in the file system. The PostgreSQL database stores only the path to these files.

A Python script was created to periodically perform the ETL process in order to clear, integrate, anonymize and load messages from SQLite to PostgreSQL database. We took into consideration privacy issues by anonymizing users' names and cell phone numbers. For this, we create an anonymous and unique ID for each user by using an MD5 hash function on its phone number. Similarly, we create an anonymous alias for each group. Since the groups are public, our approach does not violate WhatsApp's privacy policy [1].

### 3.3 Module III: Knowledge Discovery

This module explores the data stored in the PostgreSQL to finding implicit, previously unknown, and potentially useful patterns. Its main component is the Misinformation Detector, a machine learning classifier once trained and tested. This component receives a text as input and returns as output if the text is or not the misinformation. Besides, two other components are under development: a misinformation superspreader users classifier and a bot detector. It is important to highlight that the focus of this work is the design of the Digital Lighthouse platform and the orchestration of its several components. For this reason, we will not detail the algorithms, methods and strategies used in the knowledge discovery. We will do this in other papers.

### 3.4 Module IV: Data Visualization

Today, there is a great need for displaying massive amounts of data in a way that is easily accessible and understandable. In this context, data visualization is a way to represent information graphically, highlighting patterns and trends in data and helping to achieve new insights. It enables the data exploration via the manipulation of charts and images. More specifically, it enables users to analyze the data by interacting directly with a visual representation of it. In this work, the data visualization module is a web application developed using Python programming language and Django 3 framework.

## 4 CASE STUDY

To evaluate the platform proposed in this paper, we performed an exploratory case study using three different WhatsApp' messages datasets, covering relevant themes such as the Brazilian general elections campaign in 2018, the covid-19 pandemic the vaccine for covid-19. This case study was influenced by (Jedlitschka and Pfahl, 2005; Kitchenham et al., 2008; Robson and McCartan, 2016; Runeson and Höst, 2009). Then, many data analysis techniques were applied to this dataset to get insights about misinformation spread.

Next, we will describe these three datasets in detail.

- Brazilian General Elections: This dataset contains 282,601 messages, obtained from 5,364 users (cell phone chips), which participated in 59 WhatsApp public groups, in the period from August to October 2018.
- Covid-19 Pandemic: This dataset contains 228,061 messages, obtained from 10,495 users (cell phone chips), which participated in 236 WhatsApp public groups, in the period from March to June 2020.
- Vaccine for Covid-19: This dataset contains 16,056 messages, obtained from 1,857 users (cell phone chips), which participated in 175 WhatsApp public groups, in the period from December 2020 to January 2021.

Using the Data Visualization Module from the Lighthouse Platform, the user can choose a specific dataset or all data from all datasets. For simplicity, from this point onwards, all graphs will be illustrated using the Covid-19 dataset.

[1] https://www.whatsapp.com/legal/privacy-policy

## 4.1 Messages Characterization

Initially, the Lighthouse Platform shows some visualizations to characterize the used dataset. Figure 2 shows the proportion between messages with and without URL. In general, messages created to spread misinformation include URLs, often from a little-known website or blog, to give credibility. Therefore, the presence of a URL can be a criterion for selecting messages to be analyzed by fact-checkers. As you can observe in Figure 2, a significant proportion of the caught messages (9.33%) involves some URL.

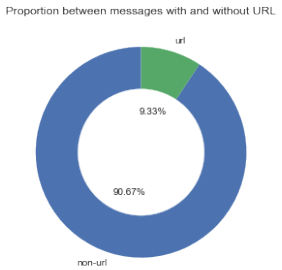


Figure 2: Proportion between Messages with and without URL.

Currently, audios, images, and videos are commonly used to spread misinformation. Therefore, the messages associated with these files are potential candidates to undergo a verification process. Figure 3 shows the proportion between messages with and without media. As you can note, 32.90% of the caught messages involves some media file.

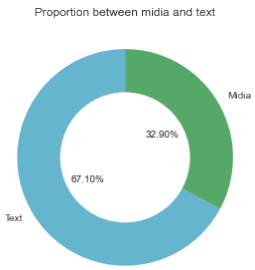


Figure 3: Proportion between Messages with and without Media.

In the 2018 Brazilian elections, many cell phone chips from foreign countries were used in the massive messaging with an electoral advertisement. Thus, monitor these messages is an important task to identify misinformation spreading. Figure 4 illustrates the proportion of foreign countries messages.

Figure 5 shows the distribution messages sending time by the day hours. As we can imagine, the peak of sending messages occurs at the time reserved for lunch (between 12 and 14 hours) and in the early evening, just after work hours.

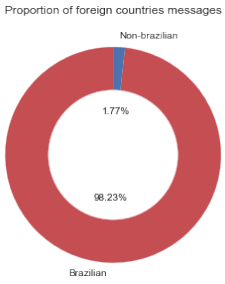


Figure 4: Proportion of Foreign Countries Messages.

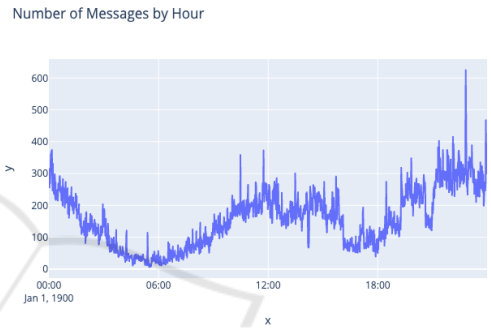


Figure 5: Number of Messages by Hour.

## 4.2 Geographic Distribution

Another relevant aspect to observe in the monitored groups is the geographic location of users (cell phone chips), both Brazilians and foreigners, besides these users' activity level. Figure 6 shows the Brazilian states with more quantity of messages. As might be expected, the most populous states have the most significant amount of messages sent. Figure 7 illustrates the Brazilian states with more users' (cell phone chips). As might be expected, the most populous states have the most significant amount of users. However, when analyzing the states with more messages per user (Figure 8), we can observe that not so populous states such as Mato Grosso do Sul, Santa Catarina, and Amazonas, have the most active users.

As previously mentioned, cell phone chips from foreign countries have been used in Brazil for massive messaging, many times spreading misinformation. Figure 9 illustrates the number of messages sent by foreign countries cell phone chips by country, while Figure 10 shows the countries with the lagers ratio between sent messages and the number of users.

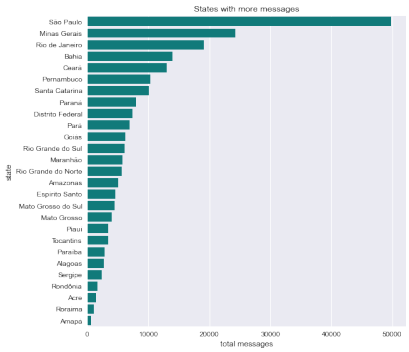


Figure 6: States with more Messages.

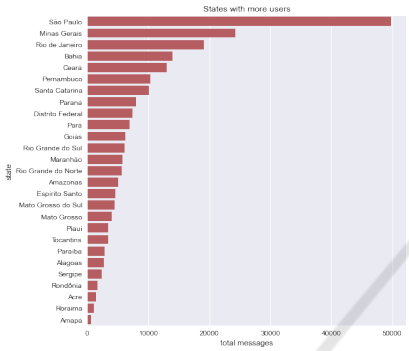


Figure 7: States with more Users.

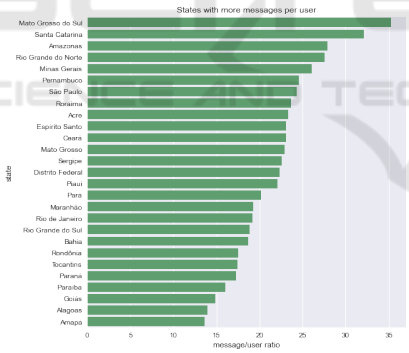


Figure 8: States with more Messages per User.

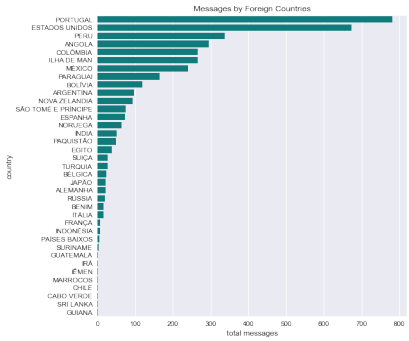


Figure 9: Messages by Foreign Countries.

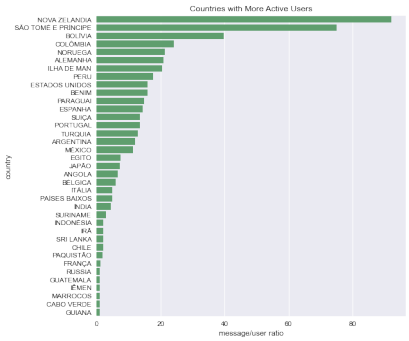


Figure 10: Countries with More Active Users.

## 4.3 Vocabulary Characterization

Another aspect that needs to be analyzed is related to the characteristics of the vocabulary used in the text messages, since there is a strong relationship between the used vocabulary and the social network, in this case, WhatsApp. Figure 11 shows the number of messages by the number of words contained in the message. As we can note, there are few messages with a large number of words and a high number of messages with few words. Figure 12 shows the word cloud highlighting the most popular words.

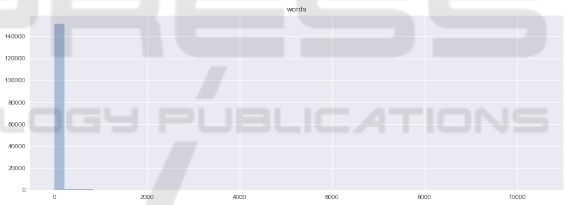


Figure 11: Number of Messages by the Number of Words in the Message.


Figure 12: Word Cloud.

## 4.4 Misinformation Analysis

The last aspect to be explored is the misinformation analysis. In this context, various information about messages and users are explored to identify text messages containing misinformation and super-spreaders.

So, we first built a new dataset adding only the messages with at least one of these words: ‘covid’, ’corona’, ’coronga’, or ’virus’. The resulting dataset had 3,014 messages. Table 1 contains the seven most shared messages. The “Sharings” column indicates how many times the message was shared. The “Mis” column indicates if the messages contains or not the misinformation. Finally, the column “NoG” denotes the number of distinct groups where the message was shared. Note that all the seven most shared messages contain misinformation.

Table 1: Most Shared Messages.

| Sharings | Text | Mis | NoG |
|---|---|---|---|
| 43 | ”PATRIOTA! *VAMOS ACORDAR BRASIL!!!! E VOCE AINDA ACREDITANDO NESTA FARSA DE COVID19, É UM GOLPE QUE FOI ARQUITETADO PARA ENGANAR OS BRASILEIROS, MENOS ESCLARECIDOS...* #NAOFIQUEEMCASA #VAMOSTRABALHAR #BOLSONAROESTACERTO | Yes | 40 |
| 26 | ”Pesquisa com mais de 6.000 médicos em 30 países diz que hidroxicloroquina é o tratamento mais eficaz para coronavírus.” | Yes | 23 |
| 23 | ”Dra. Nise Yamaguchi integra gabinete de crise e propõe a cloroquina como tratamento imediato nos casos de coronavírus.” | Yes | 23 |
| 23 | ”Herança maldita: Mandetta renova contratos de publicidade de R$ 1 bilhão firmados no governo Dilma...” | Yes | 14 |
| 22 | ”Organização Mundial de Saúde: O aborto é “essencial” durante a pandemia de coronavírus chinês.” | Yes | 22 |
| 18 | ”Prezados amigos.. vocês sabiam que, todos os problemas da humanidade foram curados com esse pânico fake do covid19?????? Vejam?? Sempre morreram milhares de pessoas de H1N1, POIS, NUNCA FOI ERRADICADA ESTA GRIPE, DE AIDS que NUNCA FOI ERRADICADA, DE TUBERCULOSE, DE INFARTO, DE BRIGAS DOMÉSTICAS, DE IDADE, DE INSUFICIÊNCIA RESPIRATÓRIA, DE CÂNCER, DE DIVERSAS OUTRAS DOENÇAS E MALES... TUDO ACABOU...” | Yes | 18 |
| 16 | ”*Atenção*: Isso a Globo não mostra. Banco Mundial acaba de lançar um documento que ressalta o papel do comércio internacional na mitigação dos impactos do coronavírus. A instituição argumenta que a manutenção dos fluxos de comércio será crucial para o suprimento de itens médicos e alimentos — e portanto limitar impactos negativos sobre empregos e nível de pobreza em escala global. O trabalho do Banco Mundial coloca o Brasil como “Exemplo 1” no quadro “Melhores Práticas em Lidar com a Covid-19”. #BolsonaroTemRazão” | Yes | 11 |

Table 2 contains the 5 most active users together with the number of messages shared by each one. The user identification was anonymized. Let’s take a particular user, for example, the user with Id -9126362355320474072, which sent 67 messages. Table 3 contains all messages shared by the user -9126362355320474072. Note that all 67 messages shared by this user have misinformation. Besides, some messages were shared many times. Now, let’s take a specific message of the user -9126362355320474072, as, for example, the message in the first row of Table 3. Table 4 contains the date and time of each sharing of the selected message, besides the group in which it was shared. Note that the selected message was shared 22 times in 22 different groups, in a period of four minutes. So, we can classify the user -9126362355320474072 as a misinformation super-spreader.

Table 2: Most Active Users.

| User Id | Number of Messages |
|---|---|
| 3346599479176653344 | 110 |
| 8121536360444460807 | 102 |
| -9126362355320474072 | 67 |
| 8900877460624761918 | 62 |
| 1721737435325801397 | 60 |

Table 3: Messages of User Id -9126362355320474072.

| Sharings | Text | Mis |
|---|---|---|
| 22 | ”Pesquisa com mais de 6.000 médicos em 30 países diz que hidroxicloroquina é o tratamento mais eficaz para coronavírus.” | Yes |
| 22 | ”Dra. Nise Yamaguchi integra gabinete de crise e propõe a cloroquina como tratamento imediato nos casos de coronavírus.” | Yes |
| 22 | ”Organização Mundial de Saúde: O aborto é “essencial” durante a pandemia de coronavírus chinês.” | Yes |
| 1 | ”ENTENDA COMO FOMOS IMPEDIDOS DE VOTAR O FUNDÃO PARA O COMBATE AO CORONAVÍRUS...” | Yes |

# 5 CONCLUSIONS

The fast spread of misinformation through WhatsApp messages poses a significant social problem. In this work, we present an entire platform, called Digital Lighthouse, that aims for finding WhatsApp public groups, besides extracting, cleaning, analyzing, and visualizing misinformation that spread in such groups. Using the proposed platform we build three different WhatsApp’ messages datasets, covering relevant themes such as the Brazilian elections, the covid-19 pandemic, and the vaccine for covid-19. Besides, we presented a case study using the pro-

Table 4: Details of the Selected Message.

| Date | Time | Group Id |
|---|---|---|
| 2020/04/06 | 18:36 | 2020_117 |
| 2020/04/06 | 18:36 | 2020_133 |
| 2020/04/06 | 18:36 | 2020_153 |
| 2020/04/06 | 18:36 | 2020_187 |
| 2020/04/06 | 18:36 | 2020_243 |
| 2020/04/06 | 18:36 | 2020_26 |
| 2020/04/06 | 18:36 | 2020_96 |
| 2020/04/06 | 18:37 | 2020_128 |
| 2020/04/06 | 18:37 | 2020_131 |
| 2020/04/06 | 18:37 | 2020_174 |
| 2020/04/06 | 18:37 | 2020_84 |
| 2020/04/06 | 18:38 | 2020_146 |
| 2020/04/06 | 18:38 | 2020_170 |
| 2020/04/06 | 18:38 | 2020_171 |
| 2020/04/06 | 18:38 | 2020_22 |
| 2020/04/06 | 18:38 | 2020_225 |
| 2020/04/06 | 18:38 | 2020_229 |
| 2020/04/06 | 18:38 | 2020_233 |
| 2020/04/06 | 18:38 | 2020_73 |
| 2020/04/06 | 18:38 | 2020_99 |
| 2020/04/06 | 18:39 | 2020_105 |
| 2020/04/06 | 18:39 | 2020_226 |

posed platform. Initially, we characterize the used dataset, explored the geographic distribution of the messages and performed a vocabulary characterization. Finally, we performed a misinformation analysis and we identified a misinformation super-spreader. As future work we will extend the Lighthouse platform using big data and real-time technologies.

## REFERENCES


Faustini, P. and Covões, T. (2019). Fake news detection using one-class classification. In *2019 8th Brazilian Conference on Intelligent Systems (BRACIS)*, pages 592–597.

Gaglani, J., Gandhi, Y., Gogate, S., and Halbe, A. (2020). Unsupervised whatsapp fake news detection using semantic search. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 285–289. IEEE.

Garimella, K. and Tyson, G. (2018). Whatsapp, doc? a first look at whatsapp public group data. *arXiv preprint arXiv:1804.01473*.

Hamdi, T., Slimi, H., Bounhas, I., and Slimani, Y. (2020). A hybrid approach for fake news detection in twitter based on user features and graph embedding. In Hung, D. V. and D'Souza, M., editors, *Distributed Computing and Internet Technology - 16th International Conference, ICDCIT 2020, Bhubaneswar, India, January 9-12, 2020, Proceedings*, volume 11969 of *Lecture Notes in Computer Science*, pages 266–280. Springer.

Jedlitschka, A. and Pfahl, D. (2005). Reporting guidelines for controlled experiments in software engineering. In *Empirical Software Engineering, 2005. 2005 International Symposium on*, pages 10–pp. IEEE.

Kitchenham, B., Al-Khilidar, H., Babar, M. A., Berry, M., Cox, K., Keung, J., Kurniawati, F., Staples, M., Zhang, H., and Zhu, L. (2008). Evaluating guidelines for reporting empirical software engineering studies. *Empirical Software Engineering*, 13(1):97–121.

Machado, C., Kira, B., Narayanan, V., Kollanyi, B., and Howard, P. (2019). A study of misinformation in whatsapp groups with a focus on the brazilian presidential elections. WWW '19, page 1013–1019, New York, NY, USA. Association for Computing Machinery.

Qiu, X., Oliveira, D. F., Shirazi, A. S., Flammini, A., and Menczer, F. (2017). Limited individual attention and online virality of low-quality information. *Nature Human Behaviour*, 1(7):0132.

Resende, G., Melo, P., Sousa, H., Messias, J., Vasconcelos, M., Almeida, J., and Benevenuto, F. (2019). (mis)information dissemination in whatsapp: Gathering, analyzing and countermeasures.

Resende, G., Messias, J., Silva, M., Almeida, J., Vasconcelos, M., and Benevenuto, F. (2018). A system for monitoring public political groups in whatsapp. In *Proceedings of the 24th Brazilian Symposium on Multimedia and the Web*, WebMedia '18, page 387–390, New York, NY, USA. Association for Computing Machinery.

Robson, C. and McCartan, K. (2016). *Real world research*. Wiley.

Runeson, P. and Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering*, 14(2):131–164.

Shu, K., Bernard, H. R., and Liu, H. (2018). Studying fake news via network analysis: Detection and mitigation. *CoRR*, abs/1804.10233.

Shu, K., Zhou, X., Wang, S., Zafarani, R., and Liu, H. (2019). The role of user profiles for fake news detection. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, ASONAM '19, page 436–439, New York, NY, USA. Association for Computing Machinery.

Silva, R. M., Santos, R. L., Almeida, T. A., and Pardo, T. A. (2020). Towards automatically filtering fake news in portuguese. *Expert Systems with Applications*, 146:113199.

Su, Q., Wan, M., Liu, X., and Huang, C.-R. (2020). Motivations, methods and metrics of misinformation detection: An nlp perspective. *Natural Language Processing Research*, 1:1–13.

Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. *Science*, 359:1146–1151.

Zhang, Y. and Hara, T. (2020). A probabilistic model for malicious user and rumor detection on social media. In *53rd Hawaii International Conference on System Sciences, HICSS 2020, Maui, Hawaii, USA, January 7-10, 2020*, pages 1–10. ScholarSpace.